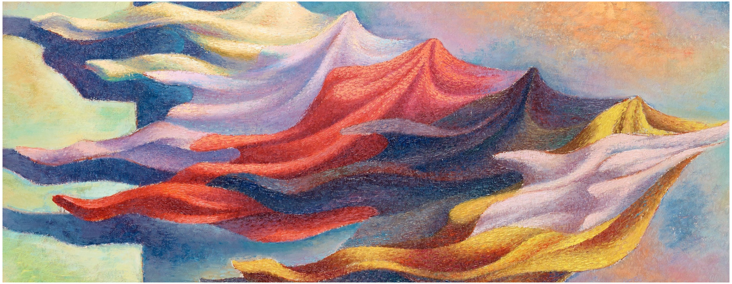


Getting the Right Tail Right

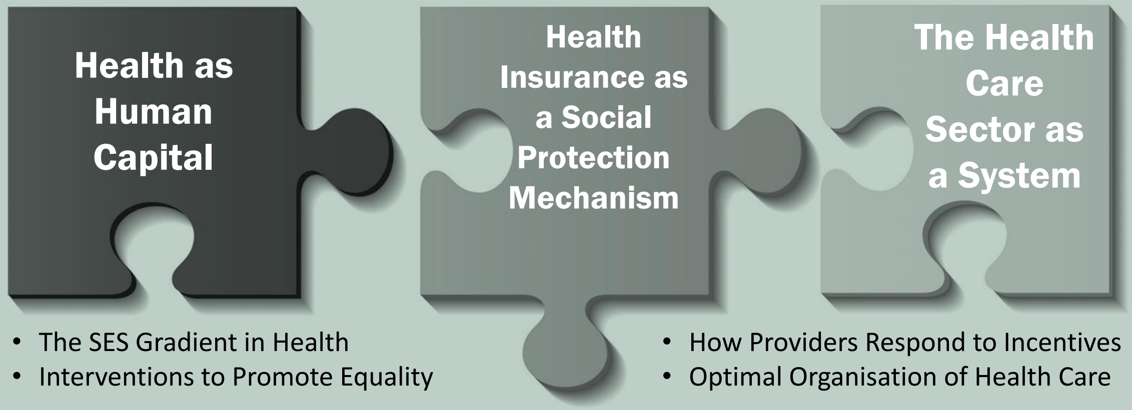
DAV/DGVFM Herbsttagung – 18 November 2025

Martin Karlsson

CINCH, University of Duisburg-Essen

Joint work with **Yulong Wang** and **Nicolas Ziebarth**.

CINCH – National Research Centre for Health Economics



- The SES Gradient in Health
- Interventions to Promote Equality

- How Providers Respond to Incentives
- Optimal Organisation of Health Care

- Risk Adjustment in Social Insurance
- Moral Hazard and Cost-Sharing

Germany: A Rising Star?

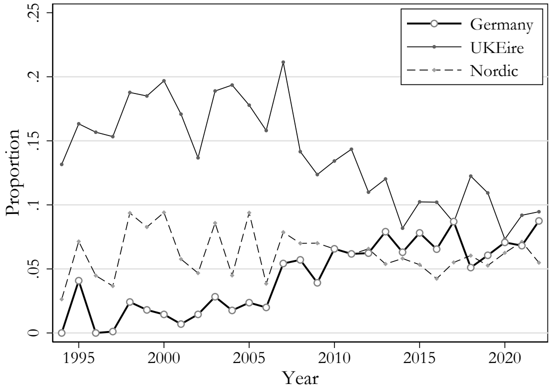


Figure 1. Proportion of Regions in Health Economics Publishing – Europe.

Source: Gschwent et al. (2025).

Diverse Research Topics

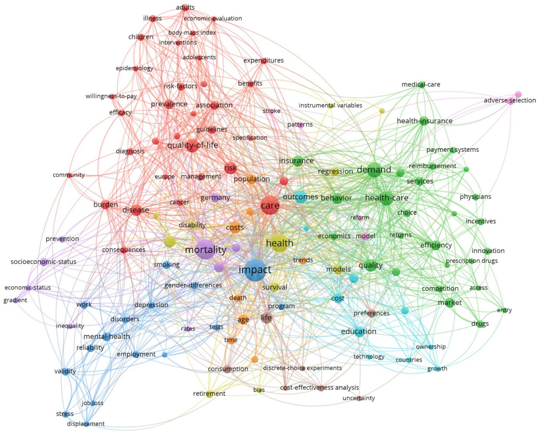


Figure 2. Health Economics Papers by Germans 1994-2022.

Source: Gschwent et al. (2025).

Introduction

The Real-World Puzzle

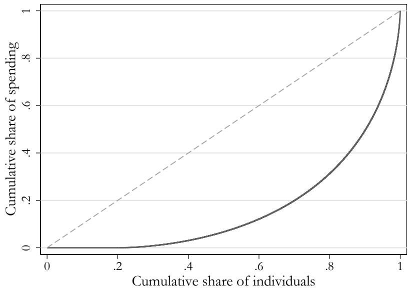


Figure 3. Lorenz Curve of PHI Expenditure.

- A small fraction of individuals account for a huge share of total health spending.

The Real-World Puzzle

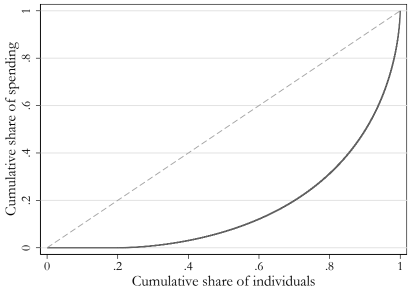


Figure 3. Lorenz Curve of PHI Expenditure.

- A small fraction of individuals account for a huge share of total health spending.
- **Example:** The top 5% of people generate about 50% of all costs.

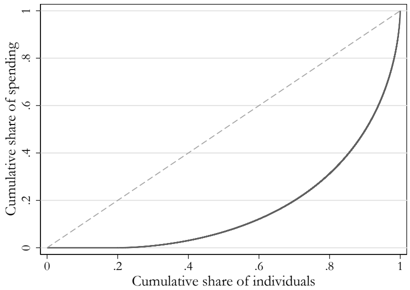


Figure 3. Lorenz Curve of PHI Expenditure.

- A small fraction of individuals account for a huge share of total health spending.
- **Example:** The top 5% of people generate about 50% of all costs.
- This concentration is well known — but what does it mean statistically?

The Real-World Puzzle

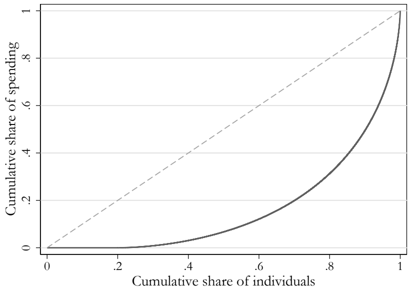


Figure 3. Lorenz Curve of PHI Expenditure.

- A small fraction of individuals account for a huge share of total health spending.
- **Example:** The top 5% of people generate about 50% of all costs.
- This concentration is well known — but what does it mean statistically?
- Can our usual regression tools handle such extreme heterogeneity?

Why This Matters for Actuaries

- Risk concentration drives **capital requirements**, **reinsurance design**, and **pricing**.
- Forecasting average costs requires models that are statistically stable.
- But when tail risks dominate, even large datasets cannot save unstable estimators.

Why This Matters for Actuaries

- Risk concentration drives **capital requirements**, **reinsurance design**, and **pricing**.
- Forecasting average costs requires models that are statistically stable.
- But when tail risks dominate, even large datasets cannot save unstable estimators.

Key idea

If the variance of health costs is infinite, regression-based risk estimates become unreliable – no matter how big your sample.

The Data

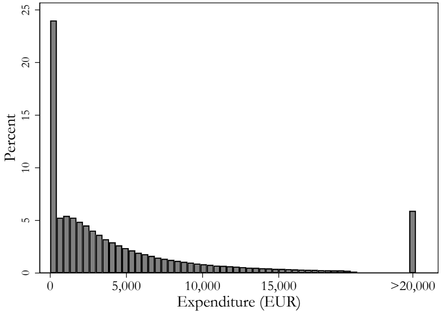


Figure 4. Histogram of Spending.

- German administrative health data: millions of enrollees, multiple years.

The Data

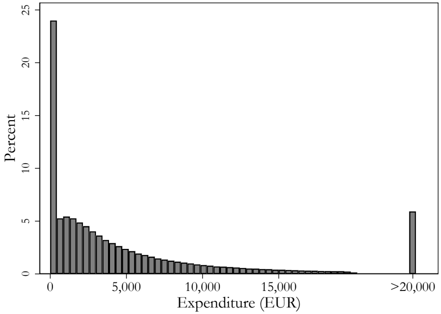


Figure 4. Histogram of Spending.

- German administrative health data: millions of enrollees, multiple years.
- Focus on individuals with full-year coverage and complete demographic info.

The Data

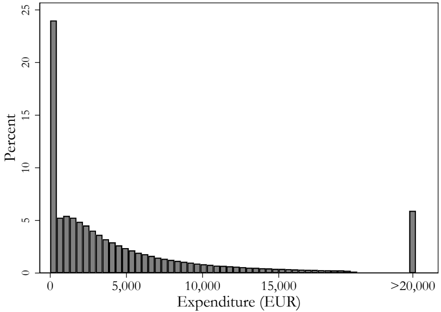


Figure 4. Histogram of Spending.

- German administrative health data: millions of enrollees, multiple years.
- Focus on individuals with full-year coverage and complete demographic info.
- Data large enough to estimate tail behaviour with confidence.

Starting Point: OLS

- Consider the canonical OLS model:

$$Y_i = X_i'\beta_0 + u_i \tag{1}$$

Starting Point: OLS

- Consider the canonical OLS model:

$$Y_i = X_i' \beta_0 + u_i \tag{1}$$

- Gauss-Markov theorem: under standard assumptions, OLS is BLUE.

Starting Point: OLS

- Consider the canonical OLS model:

$$Y_i = X_i'\beta_0 + u_i \tag{1}$$

- Gauss-Markov theorem: under standard assumptions, OLS is BLUE.
- However, $\mathbb{E}\left(Y_i^2\right) < \infty$ required for correct inference.

Starting Point: OLS

- Consider the canonical OLS model:

$$Y_i = X_i' \beta_0 + u_i \tag{1}$$

- Gauss-Markov theorem: under standard assumptions, OLS is BLUE.
- However, $\mathbb{E}\left(Y_i^2\right) < \infty$ required for correct inference.
- Typically **not** satisfied for health care expenditure.

- Compare OLS and tail-robust estimates of age gradients in expected costs.

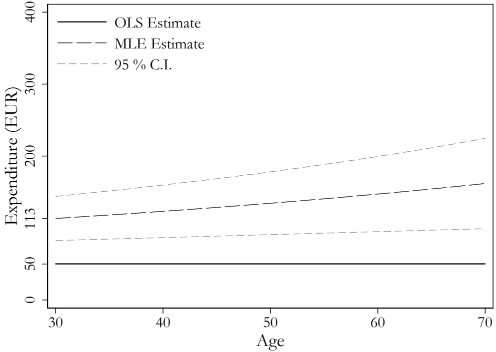


Figure 5. Comparison of OLS and MLE Estimates of Marginal Age Effects.

OLS Underestimates Age Effects When Tails Are Heavy

- Compare OLS and tail-robust estimates of age gradients in expected costs.
- OLS smooths over extremes, underestimating high-cost risk among older groups.

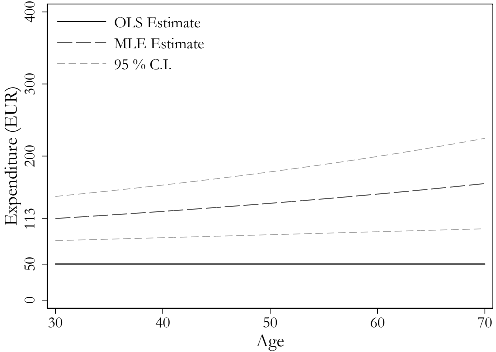


Figure 5. Comparison of OLS and MLE Estimates of Marginal Age Effects.

OLS Underestimates Age Effects When Tails Are Heavy

- Compare OLS and tail-robust estimates of age gradients in expected costs.
- OLS smooths over extremes, underestimating high-cost risk among older groups.
- Tail-aware models recover stronger, more realistic age gradients.

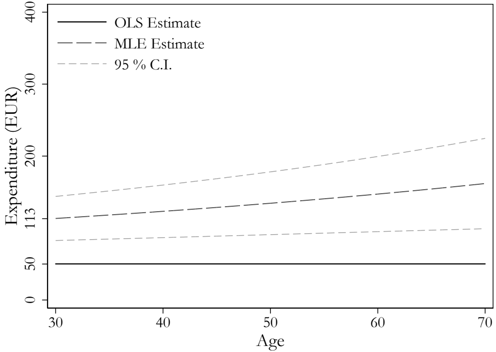


Figure 5. Comparison of OLS and MLE Estimates of Marginal Age Effects.

A Tail-Aware Alternative

- Model upper tail explicitly using a Pareto distribution above a threshold.
- Allow tail exponent α to depend on covariates (age, gender, plan type).

A Tail-Aware Alternative

- Model upper tail explicitly using a Pareto distribution above a threshold.
- Allow tail exponent α to depend on covariates (age, gender, plan type).

Key equation (simplified)

$$P(Y > y \mid X) = \left(\frac{y}{y_0}\right)^{-\alpha(X)} \tag{2}$$

Simulation Comparison

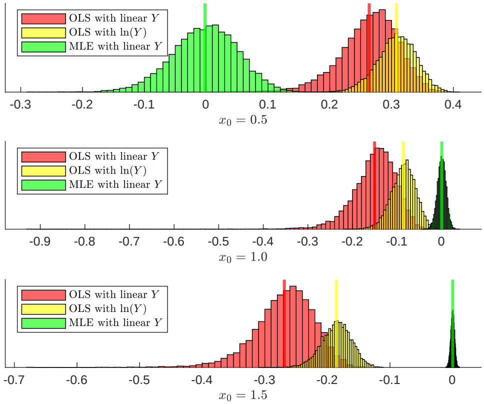


Figure 6. Histograms of OLS with Linear or Logarithm of Y and proposed MLE

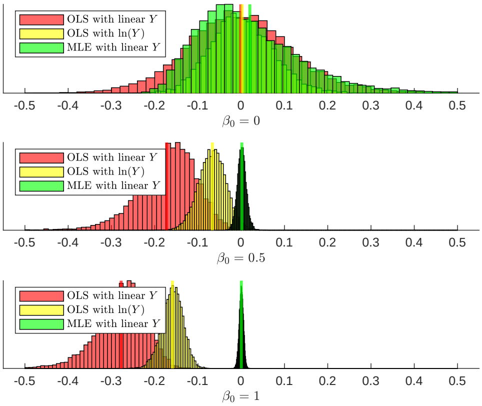


Figure 7. Histograms of OLS with Linear or Logarithm of Y and proposed MLE

Generalized Linear Models under Heavy Tails

- GLM is a standard tool in actuarial and health-economics modelling:

$$\mathbb{E}[Y_i|X_i] = \exp(\beta_0 + \beta_1 X_i)$$

with $Y_i = \exp(\beta_0 + \beta_1 X_i) + u_i$.

Generalized Linear Models under Heavy Tails

- GLM is a standard tool in actuarial and health-economics modelling:

$$\mathbb{E}[Y_i|X_i] = \exp(\beta_0 + \beta_1 X_i)$$

  with $Y_i = \exp(\beta_0 + \beta_1 X_i) + u_i$.
- Works well for skewed but light-tailed data (e.g. Gamma or lognormal).

Generalized Linear Models under Heavy Tails

- GLM is a standard tool in actuarial and health-economics modelling:

$$\mathbb{E}[Y_i|X_i] = \exp(\beta_0 + \beta_1 X_i)$$

  with $Y_i = \exp(\beta_0 + \beta_1 X_i) + u_i$.
- Works well for skewed but light-tailed data (e.g. Gamma or lognormal).
- We simulate data with heavy-tailed errors (u_i following a Generalized Pareto).

Generalized Linear Models under Heavy Tails

- GLM is a standard tool in actuarial and health-economics modelling:

$$\mathbb{E}[Y_i|X_i] = \exp(\beta_0 + \beta_1 X_i)$$

  with $Y_i = \exp(\beta_0 + \beta_1 X_i) + u_i$.
- Works well for skewed but light-tailed data (e.g. Gamma or lognormal).
- We simulate data with heavy-tailed errors (u_i following a Generalized Pareto).

Main Findings from Simulation

- For $\xi < 0.5$ ($\alpha > 2$): GLM estimates are stable and unbiased.
- For $\xi > 0.6$ ($\alpha < 2$):
  - Standard errors and confidence intervals explode.
  - t-tests overreject sharply.
  - Estimator may become numerically unstable.
- GLM assumes finite variance; heavy tails violate this.

From Two-Part to Three-Part Models

- The standard **two-part model** captures:

$$\mathbb{E}[Y|X] = \Pr(Y > 0|X) \times \mathbb{E}[Y|Y > 0, X]$$

From Two-Part to Three-Part Models

- The standard **two-part model** captures:

$$\mathbb{E}[Y|X] = \Pr(Y > 0|X) \times \mathbb{E}[Y|Y > 0, X]$$

  - Part 1: Probability of any expenditure ($Y > 0$).

From Two-Part to Three-Part Models

- The standard **two-part model** captures:

$$\mathbb{E}[Y|X] = \Pr(Y > 0|X) \times \mathbb{E}[Y|Y > 0, X]$$

  - Part 1: Probability of any expenditure ($Y > 0$).
  - Part 2: Mean of positive expenditures.

From Two-Part to Three-Part Models

- The standard **two-part model** captures:

$$\mathbb{E}[Y|X] = \Pr(Y > 0|X) \times \mathbb{E}[Y|Y > 0, X]$$

  - Part 1: Probability of any expenditure $(Y > 0)$.
  - Part 2: Mean of positive expenditures.

- Works well when expenditures are light-tailed, but:

From Two-Part to Three-Part Models

- The standard **two-part model** captures:

$$\mathbb{E}[Y|X] = \Pr(Y > 0|X) \times \mathbb{E}[Y|Y > 0, X]$$

  - Part 1: Probability of any expenditure ($Y > 0$).
  - Part 2: Mean of positive expenditures.

- Works well when expenditures are light-tailed, but:
  - A small number of extreme observations can dominate the mean.

From Two-Part to Three-Part Models

- The standard **two-part model** captures:

$$\mathbb{E}[Y|X] = \Pr(Y > 0|X) \times \mathbb{E}[Y|Y > 0, X]$$

  - Part 1: Probability of any expenditure ($Y > 0$).
  - Part 2: Mean of positive expenditures.

- Works well when expenditures are light-tailed, but:
  - A small number of extreme observations can dominate the mean.
  - GLM assumptions may fail in the tail.

From Two-Part to Three-Part Models

- The standard **two-part model** captures:

$$\mathbb{E}[Y|X] = \Pr(Y > 0|X) \times \mathbb{E}[Y|Y > 0, X]$$

  - Part 1: Probability of any expenditure $(Y > 0)$.
  - Part 2: Mean of positive expenditures.

- Works well when expenditures are light-tailed, but:
  - A small number of extreme observations can dominate the mean.
  - GLM assumptions may fail in the tail.
- **Idea:** Add a third part to handle the heavy tail explicitly.

The Three-Part Model: Capturing the Whole Distribution

- **Part 1:** Probability of being in each region ($Y = 0$, moderate, or tail) — estimated with a multinomial or sequential logit.
- **Part 2:** Conditional mean for moderate expenditures — estimated by OLS.
- **Part 3:** Conditional mean for the tail — estimated with the Pareto model.

The Three-Part Model: Capturing the Whole Distribution

- **Part 1:** Probability of being in each region ($Y = 0$, moderate, or tail) — estimated with a multinomial or sequential logit.
- **Part 2:** Conditional mean for moderate expenditures — estimated by OLS.
- **Part 3:** Conditional mean for the tail — estimated with the Pareto model.

What It Adds

- Separates drivers of **incidence**, **intensity**, and **extremes**.
- Allows for **different covariate effects** across the distribution.
- Retains interpretability and marginal-effect decomposition.

The Three-Part Model: Results

Table 1. Marginal Effects of Age: Extensive and Intensive Margins

	Pooled	Females	Males
Extensive 1	0.0041 (0.0001)	0.0011 (0.0001)	0.0060 (0.0001)
Extensive 2	0.0017 (0.0000)	0.0010 (0.0000)	0.0010 (0.0000)
Intensive 1 (OLS)	70.852 (0.63)	54.685 (1.13)	77.854 (0.74)
Intensive 2 (MLE)	141.632 (24.95)	79.207 (36.34)	202.582 (38.28)
Combined	75.380	56.079	86.468
Intensive 1+2 (OLS)	152.220 (1.68)	100.067 (2.72)	192.789 (2.13)

Practical Takeaways

- Diagnostic plots (rank–size, Lorenz) are simple tools to detect tail problems.

Practical Takeaways

- Diagnostic plots (rank–size, Lorenz) are simple tools to detect tail problems.
- If $\alpha < 2$, treat standard regressions with caution — variance is not defined.

Practical Takeaways

- Diagnostic plots (rank–size, Lorenz) are simple tools to detect tail problems.
- If $\alpha < 2$, treat standard regressions with caution — variance is not defined.
- Tail-aware models yield better forecasts of catastrophic spending.

Practical Takeaways

- Diagnostic plots (rank–size, Lorenz) are simple tools to detect tail problems.
- If $\alpha < 2$, treat standard regressions with caution — variance is not defined.
- Tail-aware models yield better forecasts of catastrophic spending.
- Important for reinsurance pricing, solvency calculations, and long-term trend analysis.

Summary and Outlook

- Extreme health costs are central to understanding health-care risk.

Summary and Outlook

- Extreme health costs are central to understanding health-care risk.
- The tail exponent and the three-part model provide a bridge between econometric analysis and actuarial practice.

Summary and Outlook

- Extreme health costs are central to understanding health-care risk.
- The tail exponent and the three-part model provide a bridge between econometric analysis and actuarial practice.
- Future work: explore tail-robust estimation in dynamic or longitudinal settings.

Summary and Outlook

- Extreme health costs are central to understanding health-care risk.
- The tail exponent and the three-part model provide a bridge between econometric analysis and actuarial practice.
- Future work: explore tail-robust estimation in dynamic or longitudinal settings.
- At CINCH, we are developing a **large-scale research programme** on these topics— and on **broader economic issues** in **German health insurance**.

Summary and Outlook

- Extreme health costs are central to understanding health-care risk.
- The tail exponent and the three-part model provide a bridge between econometric analysis and actuarial practice.
- Future work: explore tail-robust estimation in dynamic or longitudinal settings.
- At CINCH, we are developing a **large-scale research programme** on these topics— and on **broader economic issues** in **German health insurance**.
- We seek collaboration partners from the insurance sector– happy to discuss ideas and opportunities during the conference.

Thank You for Listening!

Literature I

GSCHWENT, L., B. HAMMARFELT, M. KARLSSON, AND M. KIFMANN (2025): "The Rise of Health Economics: Transforming the Landscape of Economic Research," *Health Economics*.

Diagnostics

Diagnostic: Rank Size Plots

- Characterise prevalence of extreme values by plotting $\ln(Y)$ against log rank.
- Let $Y_{(1)} \geq Y_{(2)} \geq Y_{(3)} \geq \cdots \geq Y_{(n)}$ be outcome ordered in **descending** order.
- **Linear** rank-size plot implies underlying expenditure has a **Pareto tail**.
- I.e. if Y_i has a Pareto distribution beyond some cutoff value y_{min}:

$$\Pr\left(Y_i \geq y \mid Y_i \geq y_{min}\right) = \left(\frac{y}{y_{min}}\right)^{-\alpha} \tag{3}$$

- Where α is the *Pareto exponent*.
- The slope of a linear fit in the rank-size plot identifies α.

Evidence: Rank Size Plots – Females

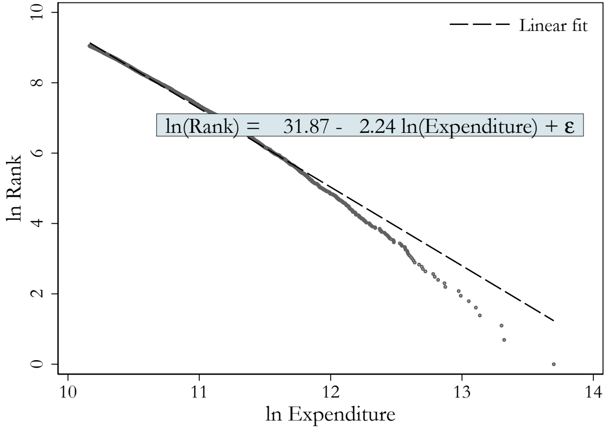


Figure 8. Rank-Size Plots of Natural Logarithms of Rank against Expenditures

Evidence: Rank Size Plots – Males

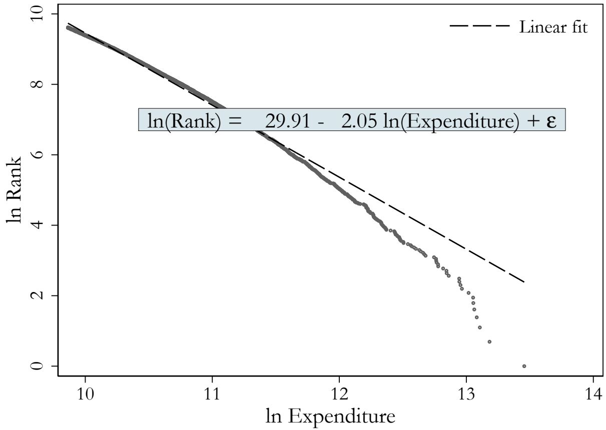


Figure 9. Rank-Size Plots of Natural Logarithms of Rank against Expenditures

Simulation Setup

Setup

- In order to assess the implications of heavy tails, we conduct a number of Monte Carlo simulations.
- We consider three different cases:
  1. **Ordinary Least Squares**:
  $$Y_i = \beta_0 + \beta_1 X_i + u_i \tag{4}$$
  with u_i comes from from a **two-sided generalized Pareto** distribution.
  2. **Generalised Linear Model**:
  $$Y_i = \exp\left(\beta_0 + \beta_1 X_i\right) + u_i \tag{5}$$
  with same distribution for u_i.
  3. **Misspecified OLS**: simulation as in (1), but estimation of
  $$\ln\left(Y_i\right) = \beta_0 + \beta_1 X_i + u_i \tag{6}$$

Evaluation

- We vary the **Pareto exponent** α between 11 and 1 – pay particular attention when it drops below 2.
- Also vary sample size between 500 and 1M.
- Performance statistics:
  1. **Mean Absolute Deviation** (MAD): $S^{-1}\sum_{s=1}^{S}|\hat{\beta}_1(s)-\beta_1|$
  2. **Root Mean Square Error** (RMSE):

$$\left(S^{-1}\sum_{s=1}^{S}|\hat{\beta}_1(s)-\beta_1|^2\right)^{1/2} \tag{7}$$

  3. **Rejection Rate**: $S^{-1}\sum_{s=1}^{S}1\left[|t(s)|>1.96\right]$
  4. Average length of **Confidence Interval**:

$$S^{-1}\sum_{s=1}^{S}2\times1.96\hat{\sigma}(s) \tag{8}$$

Simulation Results: OLS

Table 2. OLS Simulation Results with Generalized Pareto Distribution

n	500	1000	5000	10^4	10^5	10^6	500	1000	5000	10^4	10^5	10^6
$\xi(1/\alpha)$	Panel A: MAD						Panel B: RMSE					
0.09	0.06	0.04	0.02	0.01	0.00	0.00	0.07	0.05	0.02	0.02	0.01	0.00
0.19	0.07	0.05	0.02	0.02	0.01	0.00	0.09	0.06	0.03	0.02	0.01	0.00
0.29	0.09	0.06	0.03	0.02	0.01	0.00	0.12	0.08	0.04	0.03	0.01	0.00
0.39	0.13	0.09	0.04	0.03	0.01	0.00	0.18	0.12	0.05	0.04	0.01	0.00
0.49	0.18	0.14	0.07	0.05	0.02	0.01	0.27	0.25	0.12	0.09	0.02	0.01
0.59	0.32	0.24	0.13	0.10	0.05	0.02	1.31	0.68	0.25	0.16	0.08	0.03
0.69	0.63	0.50	0.28	0.23	0.12	0.06	5.09	3.36	0.89	0.79	0.47	0.49
0.79	1.14	0.94	0.74	0.64	0.38	0.24	4.47	3.98	5.43	4.63	1.84	1.40
0.89	2.87	5.19	5.02	3.43	1.88	1.19	46.9	260	291	147	30.0	13.5
0.99	5.61	6.69	5.49	5.68	5.00	7.75	44.5	87.7	44.1	38.7	37.1	156
$\xi(1/\alpha)$	Panel C: Rejection Prob.						Panel D: Length of 95% CI					
0.09	0.05	0.05	0.05	0.05	0.05	0.05	0.28	0.20	0.09	0.06	0.02	0.01
0.19	0.05	0.05	0.05	0.05	0.05	0.05	0.34	0.25	0.11	0.08	0.02	0.01
0.29	0.05	0.05	0.05	0.05	0.05	0.05	0.44	0.31	0.14	0.10	0.03	0.01
0.39	0.05	0.05	0.05	0.05	0.05	0.05	0.59	0.43	0.20	0.14	0.05	0.01
0.49	0.04	0.04	0.05	0.05	0.05	0.05	0.85	0.65	0.32	0.24	0.08	0.03
0.59	0.04	0.03	0.04	0.04	0.04	0.04	1.43	1.08	0.58	0.44	0.18	0.07
0.69	0.03	0.03	0.04	0.04	0.03	0.04	2.75	2.17	1.23	1.01	0.51	0.27
0.79	0.03	0.03	0.03	0.03	0.03	0.03	4.81	3.97	3.14	2.73	1.62	1.03
0.89	0.03	0.03	0.03	0.02	0.03	0.03	11.8	20.8	20.1	13.8	7.70	4.92
0.99	0.02	0.02	0.02	0.02	0.02	0.02	22.9	27.2	22.3	23.1	20.4	31.1

Table 3. GLM Simulation Results with Generalized Pareto Distribution

n	500	1000	5000	10^4	10^5	10^6	500	1000	5000	10^4	10^5	10^6
$\xi(1/\alpha)$	Panel A: MAD						Panel B: RMSE					
0.09	0.02	0.02	0.01	0.00	0.00	0.00	0.03	0.02	0.01	0.01	0.00	0.00
0.19	0.03	0.02	0.01	0.01	0.00	0.00	0.03	0.02	0.01	0.01	0.00	0.00
0.29	0.04	0.02	0.01	0.01	0.00	0.00	0.07	0.03	0.01	0.01	0.00	0.00
0.39	0.05	0.03	0.02	0.01	0.00	0.00	0.27	0.05	0.02	0.01	0.00	0.00
0.49	0.07	0.05	0.03	0.02	0.01	0.00	0.30	0.19	0.04	0.03	0.01	0.00
0.59	0.12	0.10	0.05	0.04	0.01	0.01	0.53	0.46	0.11	0.07	0.03	0.01
0.69	0.23	0.16	0.10	0.08	0.04	0.02	0.92	0.46	0.30	0.29	0.09	0.06
0.79	0.39	0.33	0.21	0.18	0.11	0.08	1.27	1.14	0.60	0.40	0.24	0.22
0.89	0.62	0.56	0.42	0.39	0.28	0.23	1.89	1.54	1.03	0.99	0.55	0.45
0.99	0.96	0.89	0.72	0.70	0.61	0.52	2.44	2.12	1.51	1.45	1.12	0.89
$\xi(1/\alpha)$	Panel C: Rejection Prob.						Panel D: Length of 95% CI					
0.09	0.05	0.05	0.05	0.05	0.05	0.05	0.10	0.07	0.03	0.02	0.01	0.00
0.19	0.05	0.05	0.05	0.05	0.05	0.05	0.13	0.09	0.04	0.03	0.01	0.00
0.29	0.05	0.05	0.05	0.05	0.05	0.05	0.16	0.12	0.05	0.04	0.01	0.00
0.39	0.05	0.04	0.05	0.05	0.05	0.05	0.22	0.16	0.07	0.05	0.02	0.01
0.49	0.04	0.04	0.04	0.04	0.05	0.05	0.33	1.03	0.12	0.09	0.03	0.01
0.59	0.05	0.05	0.04	0.04	0.04	0.04	66.6	0.42	0.23	0.16	0.07	0.03
0.69	0.06	0.06	0.05	0.05	0.04	0.04	$> 10^3$	$> 10^3$	409	$> 10^3$	1.94	$> 10^3$
0.79	0.08	0.08	0.07	0.07	0.05	0.04	$> 10^3$	$> 10^3$	$> 10^3$	$> 10^3$	$> 10^3$	$> 10^3$
0.89	0.10	0.11	0.11	0.11	0.09	0.08	$> 10^3$	$> 10^3$	$> 10^3$	$> 10^3$	$> 10^3$	$> 10^3$
0.99	0.13	0.13	0.14	0.14	0.13	0.14	$> 10^3$	$> 10^3$	$> 10^3$	$> 10^3$	$> 10^3$	$> 10^3$

Data

- We use data from a **large German provider** of **private health insurance**.
- Dataset includes the **universe of claims** through the 2005-11 period.
- The final sample consists of 1,867,465 enrollee-year observations from 362,783 individuals.

Descriptive Statistics

Table 4. Summary Statistics: German Claims Panel Data

	Mean	SD	Min	Max	N
Health Plan Parameters					
Total Claims (USD)	3,289	8,577	0	2,345,126	1,867,465
Annual premium (USD)	4,749	2,157	0	33,037	1,867,318
Deductible (USD)	675	659	0	3,224	1,867,465
Annual risk penalty (USD)	157	453	0	21,752	1,867,465
TOP Plan	0.377	0.485	0.0	1.0	1,867,465
PLUS Plan	0.338	0.473	0.0	1.0	1,867,465
ECO Plan	0.285	0.451	0.0	1.0	1,867,465
Socio-Demographics					
Age (in years)	45.5	11.4	25.0	99.0	1,867,465
Female	0.276	0.447	0.0	1.0	1,867,465
Policyholder since (years)	6.5	5.0	1.0	40.0	1,867,465
Client since (years)	12.8	11.0	1.0	86.0	1,867,465
Employee	0.336	0.473	0.0	1.0	1,867,465
Self-Employed	0.486	0.500	0.0	1.0	1,867,465
Civil Servant	0.132	0.338	0.0	1.0	1,867,465
Health Risk Penalty	0.358	0.480	0.0	1.0	1,867,465
Pre-Existing Condition Exempt	0.016	0.126	0.0	1.0	1,867,465